LECTURE ON

TIME SERIES DIAGNOSTIC TESTS

CHUNG-MING KUAN

Institute of Economics

Academia Sinica

This version: February 17, 2008

Address for correspondence: Institute of Economics, Academia Sinica, Taipei 115, Taiwan.

E-mail: ckuan@econ.sinica.edu.tw; URL: www.sinica.edu.tw/~ckuan

Contents

1 Introduction

Conducting various diagnostic tests is an important step in time series modeling. In the literature, there exist numerous diagnostic tests designed to examine the dependence (correlation) structure of a time series. If a time series is serially uncorrelated, no linear function of the lagged variables can account for the behavior of the current variable. If a time series is a martingale difference sequence (its precise definition will be given in Section 3.1), no function, linear or nonlinear, of the the lagged variables can characterize the behavior of the current variable. For a serially independent time series, there is no any relationship between the current and past variables. Diagnostic testing on data series thus provides information regarding how these data might be modeled. When a model is estimated, diagnostic tests can be applied to evaluate model residuals, which also serve as tests of model adequacy.

In practice, there are three classes of diagnostic tests, each focusing on a specific dependence structure of a time series. The tests of serial uncorrelatedness include the well known Q tests of Box and Pierce (1970) and Ljung and Box (1978), the robust Q^* test of Lobato, Nankervis, and Savin (2001), the spectral tests of Durlauf (1991), and the robust spectral test of Deo (2000). There are also tests of the martingale difference hypothesis, including Bierens (1982, 1984, 1990), Bierens and Ploberger (1997), Hong (1999), Dominguez and Lobato (2000), Whang (2000, 2001), Kuan and Lee (2004), and Park and Whang (2005). For the hypothesis of serial independence, two leading tests are the variance ratio test of Cochrane (1988) and the so-called BDS test of Brock, Dechert, and Scheinkman (1987); see also Campbell, Lo, and MacKinlay (1997) and Brock, Dechert, Scheinkman, and LeBaron (1996). Skaug and Tjostheim (1993), Pinkse (1998), and Hong (1999) also proposed non-parametric tests of serial independence.

There are also tests of another important time series property, namely, time reversibility. A time series is said to be time reversible if its finite-dimensional distributions are all invariant to the reversal of time indices. For example, sequences of independent random variables and stationary Gaussian ARMA processes are time reversible. On the other hand, a linear, non-Gaussian process is time irreversible in general. Tong (1990) also states that "time irreversibility is the rule rather than the exception when it comes to nonliearity" (p. 197). Thus, a test of time reversibility may be interpreted as a joint test of linearity and Gaussianity or a test of independence; see e.g., Ramsey and Rothman (1996)

and Chen, Chou, and Kuan (2000). It has been shown that a test of time reversibility is particularly powerful against asymmetric dependence (Chen and Kuan, 2002).

In this note we shall introduce some commonly used diagnostic tests for time series. We will not discuss non-parametric tests because they are, in general, not asymptotically pivotal, in the sense that their asymptotic distributions are data dependent or depend on some nuisance parameters. This note proceeds as follows. Section 2 focuses on the tests of serial uncorrelatedness. In Section 3, we discuss the tests of the martingale difference hypothesis. Section 4 presents the variance ratio test and the BDS test of serial independence. The tests of time reversibility are discussed in Section 5.

2   Tests of Serial Uncorrelatedness

Given a weakly stationary time series $\{y_t\}$, let μ denote its mean and $\gamma(\cdot)$ denote its autocovariance function, where $\gamma(i) = \mathrm{cov}(y_t, y_{t-i})$ for $i = 0, 1, 2, \ldots$. The autocorrelation function $\rho(\cdot)$ is such that $\rho(i) = \gamma(i)/\gamma(0)$. The series $\{y_t\}$ is serially uncorrelated if, and only if, its autocorrelation function is identically zero.

2.1   Q Tests

To test if $\{y_t\}$ is serially uncorrelated, existing tests of serial uncorrelatedness focus on a given number of autocorrelations and ignore $\rho(i)$ for large i. The null hypothesis is, for a given number m,

$$H_0 \colon \rho(1) = \cdots = \rho(m) = 0. \tag{2.1}$$

Let $\hat{\gamma}_T(i)$ denote the i th sample autocovariance:

$$\hat{\gamma}_T(i) = \frac{1}{T} \sum_{t=1}^{T-i} (y_t - \bar{y})(y_{t+i} - \bar{y}),$$

with $\bar{y}_T$ the sample average of y_t, and $\hat{\rho}_T(i) = \hat{\gamma}_T(i)/\hat{\gamma}_T(0)$ is the i th sample autocorrelation. For notation convenience, we shall suppress the subscript T and simply write $\bar{y}$, $\hat{\gamma}(i)$ and $\hat{\rho}(i)$.

Writing $\boldsymbol{\rho}_m = \big(\rho(1), \ldots, \rho(m)\big)'$, the null hypothesis (2.1) is $\boldsymbol{\rho}_m = \mathbf{0}$, and the estimator of $\boldsymbol{\rho}_m$ is $\hat{\boldsymbol{\rho}}_m = \big(\hat{\rho}(1), \ldots, \hat{\rho}(m)\big)'$. Under quite general conditions, it can be shown that, as

T tends to infinity,

$$\sqrt{T}(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \boldsymbol{V}),$$

where $\xrightarrow{D}$ stands for convergence in distribution, and the (i, j) th element of $\boldsymbol{V}$ is

$$v_{ij} = \frac{1}{\gamma(0)^2} \big[c_{i+1,j+1} - \rho(i)c_{1,j+1} - \rho(j)c_{1,i+1} + \rho(i)\rho(j)c_{1,1} \big],$$

with

$$c_{i+1,j+1} = \sum_{k=-\infty}^{\infty} \mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu) \big] -$$

$$\mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu) \big] \mathbb{E}\big[(y_{t+k} - \mu)(y_{t+k+j} - \mu) \big];$$

see Lobato et al. (2001).[1] Thus,

$$T(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m)' \boldsymbol{V}^{-1}(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m) \xrightarrow{D} \chi^2(m). \tag{2.2}$$

This distribution result is fundamental for the tests presented in this section.

Under the null hypothesis, $\boldsymbol{V}$ can be simplified such that $v_{ij} = c_{i+1,j+1}/\gamma(0)^2$ with

$$c_{i+1,j+1} = \sum_{k=-\infty}^{\infty} \mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu) \big].$$

In particular, when y_t are serially *independent*,

$$c_{i+1,j+1} = \begin{cases} 0, & i \neq j, \\ \gamma(0)^2, & i = j, \end{cases} \tag{2.3}$$

so that $\boldsymbol{V}$ reduces to an identity matrix. Consequently, the normalized sample autocorrelations $\sqrt{T}\hat{\rho}(i)$, $i = 1, \ldots, m$, are independent $\mathcal{N}(0, 1)$ random variables asymptotically. It follows from (2.2) that the joint test of (2.1) is

$$\mathcal{Q}_T = T\hat{\boldsymbol{\rho}}_m' \hat{\boldsymbol{\rho}}_m = T \sum_{i=1}^{m} \hat{\rho}(i)^2 \xrightarrow{D} \chi^2(m). \tag{2.4}$$

---

[1] Fuller (1976, p. 256) gives a different expressions of $\boldsymbol{V}$ in which the (i, j) th element of $\boldsymbol{V}$ is

$$v_{ij} = \sum_{k=-\infty}^{\infty} \rho(k)\rho(k - i + j) + \rho(k + j)\rho(k - i) - 2\rho(k)\rho(j)\rho(k - i) -$$

$$2\rho(k)\rho(i)\rho(k - j) + 2\rho(i)\rho(j)\rho(k)^2.$$

Note that the χ^2 limiting distribution is obtained from the condition of serial independence, which is stronger than the null hypothesis. This is the well known Q test of Box and Pierce (1970).

Remark: Basing on the result that $\sqrt{T}\hat{\rho}(i)$ are independent $\mathcal{N}(0,1)$ random variables asymptotically, many computer programs draw a confidence interval for the plot of sample autocorrelations $\hat{\rho}(i)$. For example, the 90% and 95% confidence intervals of $\hat{\rho}(i)$ are, respectively, $\pm 1.645/T^{1/2}$ and $\pm 1.96/T^{1/2}$, which permit a visual check of the significance of $\hat{\rho}(i)$.

When y_t are independent random variables with mean zero, variance σ^2, and finite 6 th moment, we have from a result of Fuller (1976, p. 242) that

$$
\mathrm{cov}\big(\sqrt{T}\hat{\rho}(i),\ \sqrt{T}\hat{\rho}(j)\big) = \begin{cases} \frac{T-i}{T} + O(T^{-1}), & i = j \neq 0, \\ O(T^{-1}), & i \neq j. \end{cases}
$$

This result provides an approximation up to $O(T^{-1})$. Then for sufficiently large T, the diagonal elements of $\boldsymbol{V}$ are approximately $(T-i)/T$, whereas the off-diagonal elements essentially vanish. This leads to the modified Q test of Ljung and Box (1978):

$$
\widetilde{\mathcal{Q}}_T = T^2 \sum_{i=1}^{m} \frac{\hat{\rho}(i)^2}{T-i} \xrightarrow{D} \chi^2(m), \tag{2.5}
$$

cf. (2.4). The Box-Pierce Q test and the Ljung-Box $\widetilde{Q}$ test are asymptotically equivalent, yet the latter ought to have better finite-sample performance because of its correction factor $(T-i)/T$. In practice, the Ljung-Box Q statistic is usually computed as

$$
T(T+2) \sum_{i=1}^{m} \frac{\hat{\rho}(i)^2}{(T-i)},
$$

which is of course asymptotically equivalent to (2.5).

Another modification of the Q test can be obtained by assuming that

$$
\mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu)\big] = 0, \tag{2.6}
$$

for each k when $i \neq j$ and for $k \neq 0$ when $i = j$. Given this assumption, $c_{i+1,j+1} = 0$ when $i \neq j$, but

$$
c_{i+1,j+1} = \mathbb{E}\big[(y_t - \mu)^2 (y_{t+i} - \mu)^2\big], \tag{2.7}
$$

when $i = j$. Thus, $\boldsymbol{V}$ is diagonal with the diagonal element $v_{ii} = c_{i+1,i+1}/\gamma(0)^2$, which can be consistently estimated by

$$\hat{v}_{ii} = \frac{\frac{1}{T}\sum_{t=1}^{T-i}(y_t - \bar{y})^2(y_{t+i} - \bar{y})^2}{[\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2]^2}.$$

Then under the null hypothesis, we have the following test due to Lobato et al. (2001):

$$\mathcal{Q}_T^* = T\sum_{i=1}^{m} \hat{\rho}(i)^2/\hat{v}_{ii} \xrightarrow{D} \chi^2(m). \tag{2.8}$$

Note that the Q^* test does not require serial independence of y_t, in contrast with the Box-Pierce and Ljung-Box Q tests, but it requires (2.6) which is not easy to interpret.

It can be seen that (2.7) would reduce to (2.3) if y_t are conditionally homoskedastic. Thus, estimating $c_{i+1,j+1}$ makes the Q^* test more robust to conditional heteroskedasticity, such as ARCH (autoregressive conditional heteroskedasticity) and GARCH (generalized ARCH) processes. Note that a process may be serially uncorrelated yet conditional heteroskedasitic. The Q^* test is to be preferred in practice because of its robustness.

Remarks:

1. When the Q-type tests are applied to the residuals of an ARMA(p,q) model, the asymptotic null distribution becomes $\chi^2(m - p - q)$.

2. The asymptotic distribution of the Box-Pierce and Ljung-Box Q tests is derived under the assumption that $\{y_t\}$ is serially independent. This distribution result is also valid when $\{y_t\}$ is a martingale difference sequence with additional moment conditions. Thus, these Q tests can also be interpreted as tests of independence (or martingale difference), with a focus on autocorrelations.

3. The asymptotic null distribution of the Q-type tests is valid provided that data possess at least finite $(4 + \delta)$ th moment for some $\delta > 0$. Many financial time series, unfortunately, may not satisfy this moment requirement; see e.g., de Lima (1997). For such time series, the limiting χ^2 distribution of the Q tests may not be a good approximation to its finite-sample counterpart.

2.2   The Spectral Tests

Instead of testing a fixed number of autocorrelations $\rho(j)$, it is also possible to asymptotically test if $\rho(j)$ are all zero:

$$H_0\colon \rho(1) = \rho(2) = \cdots = 0. \tag{2.9}$$

Recall that the spectral density function is the Fourier transform of the autocorrelations:

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) e^{-ij\omega}, \qquad \omega \in [-\pi, \pi],$$

where $i = (-1)^{1/2}$ and ω is the frequency. When the autocorrelations are all zero, the spectral density reduces to the constant $(2\pi)^{-1}$ for all ω. It is then natural to base a test of all autocorrelations by comparing the sample counterpart of $f(\omega)$ and $(2\pi)^{-1}$.

Let $I_T(\omega)$ denote the *periodogram*, the sample spectral density, of the time series $\{y_t\}$. The difference between $I_T(\omega)$ and $(2\pi)^{-1}$ is

$$\frac{1}{2\pi} \left(\sum_{j=-(T-1)}^{T-1} \hat{\rho}(j) e^{-ij\omega} - 1 \right),$$

which should be "close" to zero for all ω under the null hypothesis. Recall that $\exp(-ij\omega) = \cos(j\omega) - i\sin(j\omega)$, where sin is an odd function such that $\sin(j\omega) = -\sin(-j\omega)$, and cos is an even function such that $\cos(j\omega) = \cos(-j\omega)$. Then,

$$\frac{1}{2\pi} \left(\sum_{j=-(T-1)}^{T-1} \hat{\rho}(j) e^{-ij\omega} - 1 \right) = \frac{1}{\pi} \sum_{j=1}^{T-1} \hat{\rho}(j) \cos(j\omega).$$

Integrating this function with respect to ω on $[0, a]$, $0 \le a \le \pi$, we obtain the cumulated differences:

$$\frac{1}{\pi} \sum_{j=1}^{T-1} \hat{\rho}(j) \frac{\sin(ja)}{j},$$

which should also be "close" to zero for all a under the null hypothesis. The spectral test of Durlauf (1991) is then based on the normalized, cumulated differences:

$$D_T(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \hat{\rho}(j) \frac{\sin(j\pi t)}{j}, \tag{2.10}$$

where $\pi t = a$ and $m(T)$ is less than T but grows with T at a slower rate.

Recall that a standard Brownian motion B can be approximated by a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables $\{\epsilon_t\}$ via the following expression:

$$W_T(t) = \epsilon_0 t + \frac{\sqrt{2}}{\pi} \sum_{j=1}^{T} \epsilon_j \frac{\sin(j\pi t)}{j} \Rightarrow B(t), \quad t \in [0,1],$$

where $\Rightarrow$ stands for weak convergence. Then,

$$W_T(t) - tW_T(1) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{T} \epsilon_j \frac{\sin(j\pi t)}{j} \Rightarrow B^0(t), \quad t \in [0,1], \tag{2.11}$$

where B^0 denotes the Brownian bridge. It is readily seen that (2.10) is quite similar to the right-hand side of (2.11).

Recall that, under the null hypothesis, $T^{1/2}\hat{\rho}(j)$ converge in distribution to independent $\mathcal{N}(0,1)$ random variables when (i) y_t are serially independent, or (ii) y_t satisfy (2.6) with conditional homoeskedasticity. Basing on the approximation (2.11) and the asymptotic normality of $T^{1/2}\hat{\rho}(j)$, we have from (2.10) that

$$D_T(t) \Rightarrow B^0(t), \qquad t \in [0,1]. \tag{2.12}$$

The spectral tests of Durlauf (1991) are based on various functionals on D_T.

Specifically, Durlauf (1991) considered the following test statistics whose limits follow easily from (2.12) and the continuous mapping theorem.

(1) Anderson-Darling test:

$$\mathrm{AD}_T = \int_0^1 \frac{[D_T(t)]^2}{t(1-t)} \, \mathrm{d}t \Rightarrow \int_0^1 \frac{[B^0(t)]^2}{t(1-t)} \, \mathrm{d}t;$$

(2) Cramér-von Mises test:

$$\mathrm{CVM}_T = \int_0^1 [D_T(t)]^2 \, \mathrm{d}t \Rightarrow \int_0^1 [B^0(t)]^2 \, \mathrm{d}t;$$

(3) Kolmogorov-Smirnov test:

$$\mathrm{KS}_T = \sup |D_T(t)| \Rightarrow \sup |B^0(t)|;$$

(4) Kuiper test:

$$\mathrm{Ku}_T = \sup_{s,t} |D_T(t) - D_T(s)| \Rightarrow \sup |B^0(t) - B^0(s)|.$$

These limits are also the limits of the well-known goodness-of-fit tests in the statistics literature. Some of their critical values have been tabulated in, e.g., Shorack and Wellner (1986).

Similar to the finding of Lobato et al. (2001), Deo (2000) noted that when y_t are conditionally heteroskedastic, the asymptotic variance of $T^{1/2}\hat{\rho}(j)$ would be $\mathbb{E}(y_t^2 y_{t-j}^2)/\gamma(0)^2$. In this case, (2.12) fails to hold because D_T is not properly normalized, and hence D_T converges to a different weak limit. As a result, the limiting distributions of the tests considered by Durlauf (1991) have thicker right-tails under conditional heteroskedasticity and render these tests over-sized. That is, these tests reject too often under the null hypothesis than they should.

Deo (2000) proposed the following modification of D_T in (2.10):

$$D_T^c(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \frac{\hat{\rho}(j)}{\sqrt{\hat{v}_{jj}}} \frac{\sin(j\pi t)}{j},$$

where

$$\sqrt{\hat{v}_{jj}} = \frac{1}{\hat{\gamma}(0)} \left(\frac{1}{T-j} \sum_{t=1}^{T-j} (y_t - \bar{y})^2 (y_{t+j} - \bar{y})^2 \right)^{1/2}.$$

With additional regularity conditions, Deo (2000) consider the modified Cramér-von Mises test based on D_T^c:

$$\mathrm{CVM}_T^c = \int_0^1 [D_T^c(t)]^2 \, \mathrm{d}t \Rightarrow \int_0^1 [B^0(t)]^2 \, \mathrm{d}t.$$

This modification is analogous to the Q^* test of Lobato et al. (2001). The simulation results of Deo (2000) demonstrate that the modified spectral test is indeed robust to some series that are conditionally heteroskedastic.

Remarks:

1. Durlauf (1991) refers to the spectral test as a test of the martingale difference hypothesis. It should be noted, however, that the condition of *all* autocorrelations being zero is necessary but not sufficient for the martingale difference hypothesis.

2. The Durlauf test compares the periodogram with the spectral density under the null, i.e., $(2\pi)^{-1}$. Other nonparametric tests may be constructed by comparing a nonparametric estimate of $f(\omega)$ with $(2\pi)^{-1}$.

3 Tests of Martingale Difference

3.1 Martingale Difference

Let $\{\mathcal{F}^t\}$ denote a sequence of information sets such that $\mathcal{F}^s \subseteq \mathcal{F}^t$ for all $s < t$. Such a sequence is known as a *filtration*. A sequence of integrable random variables $\{y_t\}$ is said to be a *martingale difference sequence* with respect to the filtration $\{\mathcal{F}^t\}$ if, and only if, $\mathbb{E}(y_t \mid \mathcal{F}^{t-1}) = 0$ for all t. When $\{y_t\}$ is a martingale difference sequence, its cumulated sums, $\eta_t = \sum_{s=1}^t y_s$, are such that $\mathbb{E}(\eta_t \mid \mathcal{F}^{t-1}) = \eta_{t-1}$ and form the process known as a *martingale*. By the law of iterated expectations, a martingale difference sequence must have mean zero. This implication is not restrictive, as we can always evaluate the "centered" series $\{y_t - \mathbb{E}(y_t)\}$ when y_t have non-zero means.

The concept of martingale difference can be related to time series *non-predictability*. We say that $\{y_t\}$ is not predictable with respect to the filtration $\{\mathcal{F}^t\}$ (in the mean-squared-error sense) if, and only if, the conditional expectations $\mathbb{E}(y_t \mid \mathcal{F}^{t-1})$ are the same as the unconditional expectations $\mathbb{E}(y_t)$. That is, the conditioning variables in $\mathcal{F}^{t-1}$ do not help to improve on the forecast of y_t, so that the best L_2 forecast is not different from the naive forecast. Clearly, this definition is equivalent to requiring $\{y_t - \mathbb{E}(y_t)\}$ being a martingale difference sequence with respect to $\{\mathcal{F}^t\}$.

In the time series analysis, it is quite common to choose the information sets $\mathcal{F}^t$ as the σ-algebras generated by $\boldsymbol{Y}^t = \{y_t, y_{t-1}, \ldots, y_1\}$. The martingale difference property in this context is thus equivalent to the predictability of y_t based on its past information $\boldsymbol{Y}^{t-1}$. It should be noted that the predictability defined in this way is very restrictive, because y_t may not be predictable based on $\boldsymbol{Y}^{t-1}$ but may become predictable when the information sets are expanded.

It is well known that $\{y_t\}$ is a martingale difference sequence if, and only if, y_t are uncorrelated with $h(\boldsymbol{Y}^{t-1})$ for any measurable function h, i.e.,

$$\mathbb{E}[y_t\, h(\boldsymbol{Y}^{t-1})] = 0, \qquad \forall h. \tag{3.1}$$

Taking h in (3.1) as the linear function, we immediately see that y_t must be serially uncorrelated with $y_{t-1}, \ldots, y_1$. Thus, a martingale difference sequence must be serially uncorrelated; the converse need not be true. For example, consider the following nonlinear

moving average process:

$$y_t = \varepsilon_{t-1}\varepsilon_{t-2}(\varepsilon_{t-2} + \varepsilon_t + 1),$$

where ε_t are i.i.d. $\mathcal{N}(0,1)$ random variables. It is clear that $\mathrm{corr}(y_t, y_{t-j}) = 0$ for all j, yet $\{y_t\}$ is not a martingale difference process. Recall that a white noise is a sequence of uncorrelated random variables that have zero mean and a constant variance. Thus, a martingale difference sequence need not be a white noise because the former does not have any restriction on variance or other high-order moments.

If $\{y_t\}$ is a sequence of serially independent random variables with zero mean, we have

$$\mathbb{E}[y_t\, h(\boldsymbol{Y}^{t-1})] = \mathbb{E}(y_t)\,\mathbb{E}[h(\boldsymbol{Y}^{t-1})] = 0,$$

for all measurable functions h. Thus, $\{y_t\}$ is necessarily a martingale difference sequence, but the converse need not be true. Consider a simple ARCH process $\{y_t\}$ such that $y_t = v_t^{1/2}\,\varepsilon_t$, where ε_t are i.i.d. random variables with mean zero and variance 1, and

$$v_t = a + by_{t-1}^2,$$

with $a, b > 0$. While $\{y_t\}$ is a martingale difference sequence, it is serially dependent due to the correlations among y_t^2. We thus conclude that serial independence implies martingale difference which in turn implies serial uncorrelatedness.

3.2   Tests of Martingale Difference

We have seen that the martingale difference hypothesis is equivalent to (3.1). Testing this hypothesis is not easy because it would be practically infeasible to test against all measurable functions. A well known approach is to consider a class of functions, indexed by a (nuisance) parameter, that are capable of spanning the space of functions on conditioning variables. This approach is pioneered by Bierens (1982, 1984, 1990) and is based on the exponential function, yet it leads to a test that is not asymptotically pivotal; see also Bierens and Ploberger (1997). Kuan and Lee (2004) integrate out the the nuisance parameter in the exponential function and obtain a new set of testing functions. Their tests have a standard limiting distribution, but they check only a necessary condition of the martingale difference hypothesis. Park and Whang (2005) base their test on the indicator function. There are also various tests based on nonparametric approximations; see, e.g., Hong (1999).

4 Tests of Serial Independence

Testing serial independence is even more challenging because it is required to evaluate all possible relations between the variable of interest and its lagged variables (or variables in the conditioning set). In practice, diagnostic tests of serial independence typically focus on certain aspects of the data, such as serial correlations or ARCH-type dependence (i.e., squared correlations). For example, McLeod and Li (1983) suggest testing whether the first m autocorrelations of y_t^2 are zero using a Q test. That is, one computes (2.5) or its variant where $\hat{\rho}(i)$ are the sample autocorrelations of y_t^2:

$$\hat{\rho}(i) = \frac{\frac{1}{T}\sum_{t=1}^{T-i}(y_t^2 - m_2)(y_{t+i}^2 - m_2)}{\frac{1}{T}\sum_{t=1}^{T}(y_t^2 - m_2)^2},$$

with m_2 the sample mean of y_t^2. Under stronger conditions, the asymptotic null distribution of the resulting Q test remains $\chi^2(m)$. Compared with other tests of serial uncorrelatedness, this test checks another necessary condition of independence. In this section, we shall introduce some commonly used tests of i.i.d. condition which is sufficient for serial independence.

4.1 The Variance Ratio Test

The variance-ratio test of Cochrane (1988) is a convenient diagnostic test of the i.i.d. hypothesis. Consider random variables y_t. If these variables are i.i.d., $\eta_t = \sum_{i=0}^{t} y_i$ form a random walk. Thus, the variance-ratio test is also used in applications to check the random walk hypothesis.

Suppose that y_t have mean zero and variance σ^2, with $t = 0, 1, \ldots, kT$ for some k and T. Let $\hat{\sigma}^2$ be the sample variance of y_t and $\tilde{\sigma}_k^2$ an estimator of $\mathrm{var}(y_t + \cdots + y_{t-k+1})$ which is $k\sigma^2$ under the null hypothesis. As a result, $\tilde{\sigma}_k^2/k$ and $\hat{\sigma}^2$ should be close to each other under the null. The variance ratio test is simply based on a normalized version of $\tilde{\sigma}_k^2/(k\hat{\sigma}^2)$. Define the sample average of $y_1, \ldots, y_{kT}$ as

$$\bar{y} = \frac{1}{kT}\sum_{t=1}^{kT}(\eta_t - \eta_{t-1}) = \frac{1}{kT}(\eta_{kT} - \eta_0).$$

The standard variance estimator of $\sigma^2 = \mathrm{var}(\eta_t - \eta_{t-1})$ is

$$\hat{\sigma}^2 = \frac{1}{kT}\sum_{t=1}^{kT}\left(\eta_t - \eta_{t-1} - \bar{y}\right)^2.$$

Under the i.i.d. null hypothesis, it can be shown that $\sqrt{kT}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{D} \mathcal{N}(0, 2\sigma^4)$. Consider the following estimator of $\sigma_k^2 = \text{var}(\eta_t - \eta_{t-k})$:

$$\tilde{\sigma}_k^2 = \frac{1}{T}\sum_{t=1}^{T}(\eta_{kt} - \eta_{kt-k} - k\bar{y})^2 = \frac{1}{T}\sum_{t=1}^{T}\big[k(\bar{y}_t - \bar{y})\big]^2,$$

where $\bar{y}_t = \sum_{kt-k+1}^{kt} y_i/k$. Note that there are k observations for each t. Under the i.i.d. hypothesis, $\sigma_k^2 = k\sigma^2$, so that $\sqrt{T}(\tilde{\sigma}_k^2 - k\sigma^2) \xrightarrow{D} \mathcal{N}(0, 2k^2\sigma^4)$.

While $\hat{\sigma}^2$ is both consistent and asymptotically efficient for σ^2 under the null hypothesis, $\tilde{\sigma}_k^2/k$ is consistent but not asymptotically efficient. Writing

$$\frac{1}{\sqrt{k}}\sqrt{T}(\tilde{\sigma}_k^2 - k\sigma^2) = \sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \sigma^2\right)$$
$$= \sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \hat{\sigma}^2\right) + \sqrt{kT}(\hat{\sigma}^2 - \sigma^2), \tag{4.1}$$

the two terms on the right-hand side must be asymptotically uncorrelated by Hausman (1978).[2] The fact that the left-hand side of (4.1) is distributed as $\mathcal{N}(0, 2k\sigma^4)$ thus implies

$$\sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \hat{\sigma}^2\right) \xrightarrow{D} \mathcal{N}\big(0, 2(k-1)\sigma^4\big).$$

It follows that

$$\sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k\hat{\sigma}^2} - 1\right) \xrightarrow{D} \mathcal{N}\big(0, 2(k-1)\big).$$

Denoting the ratio $\tilde{\sigma}_k^2/(k\hat{\sigma}^2)$ as $\text{VR}(k)$, we obtain the variance ratio test:

$$\sqrt{kT}[\text{VR}(k) - 1]/\sqrt{2(k-1)} \xrightarrow{D} \mathcal{N}(0, 1),$$

under the null hypothesis that the data are i.i.d. It is clear that this tests depends on k.

In practice, one may employ other estimators for the variance ratio test. For example, σ_k^2 may be estimated by

$$\tilde{\sigma}_k^2 = \frac{1}{kT}\sum_{t=k}^{kT}(\eta_t - \eta_{t-k} - k\bar{y})^2.$$

---

[2]Let $\hat{\theta}_e$ be a consistent and asymptotically efficient estimator of the parameter θ and $\hat{\theta}_c$ a consistent estimator but not asymptotically efficient. Writing $\hat{\theta}_c = \hat{\theta}_c - \hat{\theta}_e + \hat{\theta}_e$, Hausman (1978) showed that $\hat{\theta}_e$ is asymptotically uncorrelated with $\hat{\theta}_c - \hat{\theta}_e$. For if not, there would exist a linear combination of $\hat{\theta}_e$ and $\hat{\theta}_c - \hat{\theta}_e$ that is asymptotically more efficient than $\hat{\theta}_e$.

One may also correct the bias of variance estimators and compute

$$\hat{\sigma}^2 = \frac{1}{kT-1} \sum_{t=1}^{kT} (\eta_t - \eta_{t-1} - \bar{y})^2,$$

$$\tilde{\sigma}_k^2 = \frac{1}{M} \sum_{t=k}^{kT} (\eta_t - \eta_{t-k} - k\bar{y})^2,$$

where $M = (kT - k + 1)(1 - 1/T)$. See Taylor (2005) for the test with a different variance estimator; Campbell, Lo, and MacKinlay (1997) provide detailed discussion of this test.

4.2   The BDS Test

The BDS test of serial independence also checks whether a sequence of random variables are i.i.d. Let $\boldsymbol{Y}_t^n = (y_t, y_{t+1}, \ldots, y_{t+n-1})$. Define the *correlation integral* with the *dimension* n and *distance* ϵ as:

$$C(n, \epsilon) = \lim_{T \to \infty} \binom{T-n}{2}^{-1} \sum_{\forall s < t} \mathbf{I}_\epsilon(\boldsymbol{Y}_t^n, \boldsymbol{Y}_s^n),$$

where $\mathbf{I}_\epsilon(Y_t^n, Y_s^n) = 1$ if the maximal norm $\|\boldsymbol{Y}_t^n - \boldsymbol{Y}_s^n\| < \epsilon$ and 0 otherwise. We may also define

$$\mathbf{I}_\epsilon(Y_t^n, Y_s^n) = \prod_{i=0}^{n-1} \mathbf{1}_{\{|y_{t+i} - y_{s+i}| < \epsilon\}},$$

where $\mathbf{1}_A$ is the indicator function of the event A. The correlation integral is a measure of the proportion that any pairs of n-vectors ($\boldsymbol{Y}_t^n$ and $\boldsymbol{Y}_s^n$) are within a certain distance ϵ.

If y_t are indeed i.i.d., $\boldsymbol{Y}_t^n$ should exhibit no pattern in the n-dimensional space, so that $C(n, \epsilon) = C(1, \epsilon)^n$. The BDS test is designed to check whether the sample counterparts of $C(n, \epsilon)$ and $C(1, \epsilon)^n$ are sufficiently close. Specifically, the BDS statistic reads

$$B_T(n, \epsilon) = \sqrt{T - n + 1}(C_T(n, \epsilon) - C_T(1, \epsilon)^n)/\hat{\sigma}(n, \epsilon),$$

where

$$C_T(n, \epsilon) = \binom{T-n}{2}^{-1} \sum_{\forall s < t} \mathbf{I}_\epsilon(\boldsymbol{Y}_t^n, \boldsymbol{Y}_s^n),$$

and $\hat{\sigma}^2(n, \epsilon)$ is a consistent estimator of the asymptotic variance of $\sqrt{T - n + 1}C_T(n, \epsilon)$; see Brock et al. (1996) for details. It has been shown that the asymptotic null distribution of the BDS test is $\mathcal{N}(0, 1)$.

The performance of the BDS test depends on the choice of n and ϵ. There is, however, no criterion to determine these two parameters. In practice, one may consider several values of n and set ϵ as a proportion to the sample standard deviation s_T of the data, i.e., $\epsilon = \delta s_T$ for some δ. Common choices of δ are 0.75, 1, and 1.5. An advantage of the BDS test is that, owing to the indicator function in the statistic, it is robust to random variables that do not possess high-order moments. The BDS test usually needs a large sample to ensure proper performance. Moreover, it has been found that the BDS test has low power against certain forms of nonlinearity such as self-exciting threshold AR processes (Rothman, 1992) and neglected asymmetry in volatility (Hsieh, 1991; Brooks and Henry, 2000; Chen and Kuan, 2002); see also Hsieh (1989, 1993) and Brooks and Heravi (1999).

5 Tests of Time Reversibility

A different type of diagnostic test focuses on the property of *time reversibility*. A strictly stationary process $\{y_t\}$ is said to be time reversible if its finite-dimensional distributions are all invariant to the reversal of time indices. That is,

$$F_{t_1, t_2, \ldots, t_n}(c_1, c_2, \ldots, c_n) = F_{t_n, t_{n-1}, \ldots, t_1}(c_1, c_2, \ldots, c_n).$$

When this condition does not hold, $\{y_t\}$ is said to be *time irreversible*. Clearly, independent sequences and stationary Gaussian ARMA processes are time reversible. Rejecting the null hypothesis of time reversibility thus implies that the data can not be serially independent. As such, the test of time reversibility can also be interpreted as a test of serial independence.

Time irreversibility indicates some time series characteristics that can not be described by the autocorrelation function. When $\{y_t\}$ is time reversible, it can be shown that for any k, the marginal distribution of $y_t - y_{t-k}$ must be symmetric about the the origin. To see this, let $A(x) = \{(a, b) : b - a \leq x\}$ and $B(x) = \{(a, b) : b - a \geq -x\}$. Then,

$$\int_{B(x)} dF_{t, t-k}(a, b) = 1 - \int_{A(-x)} dF_{t, t-k}(a, b),$$

As time reversibility implies $F_{t, t-k}(a, b) = F_{t, t-k}(b, a)$, we have

$$\int_{B(x)} dF_{t, t-k}(a, b) = \int_{B(x)} dF_{t, t-k}(b, a) = \int_{A(x)} dF_{t, t-k}(b, a) = \int_{A(x)} dF_{t, t-k}(a, b).$$

This shows that

$$\int_{A(x)} dF_{t,t-k}(a,b) = 1 - \int_{A(-x)} dF_{t,t-k}(a,b),$$

so that $y_t - y_{t-k}$ has a symmetric distribution for each k. See also Cox (1981) and Chen, Chou, and Kuan (2000).

If this symmetry fails for some k, there is some asymmetric dependence between y_t and y_{t-k}, in the sense that the effect of y_{t-k} on y_t is different from that of y_t on y_{t-k}. In view of this symmetry property, we may infer that nonlinear time series are time irreversible in general. Moreover, linear and stationary processes with non-Gaussian innovations are also typically time irreversible. Compared with existing tests of serial independence, a test of time reversibility has a different focus on time series properties and thus serves as a useful diagnostic test.

Existing tests of time reversibility aim at checking symmetry of $y_t - y_{t-k}$ for each k. A necessary condition of distribution symmetry is its third central moment being zero. One may then test time reversibility by evaluating whether the sample third moment is sufficiently close to zero. Observe that by stationarity,

$$\mathbb{E}(y_t - y_{t-k})^3 = \mathbb{E}(y_t^3) - 3\,\mathbb{E}(y_t^2 y_{t-k}) + 3\,\mathbb{E}(y_t y_{t-k}^2) - \mathbb{E}(y_{t-k}^3)$$

$$= -3\,\mathbb{E}(y_t^2 y_{t-k}) + 3\,\mathbb{E}(y_t y_{t-k}^2),$$

where the two terms on the right-hand side are referred to as the *bi-covariances*. Ramsey and Rothman (1996) base their test of time reversibility on the sample bi-covariances. Note that both the third-moment test and bi-covariance test require the data to possess at least finite 6th moment. Unfortunately, most financial time series do not satisfy this moment condition. On the other hand, Chen, Chou and Kuan (2000) consider a different testing approach that is robust to the failure of moment conditions.

It is well known that a distribution is symmetric if, and only if, the imaginary part of its characteristic function is zero. Hence, time reversibility of $\{y_t\}$ implies that

$$h_k(\omega) := \mathbb{E}\big[\sin\big(\omega(y_t - y_{t-k})\big)\big] = 0, \qquad \text{for all } \omega \in \mathbb{R}^+. \tag{5.1}$$

Note that (5.1) include infinitely moment conditions indexed by ω. Let g be a positive function such that $\int g(\omega)\,d\omega < \infty$. By changing the orders of integration, (5.1) implies that

$$\int_{\mathbb{R}^+} h_k(\omega)g(\omega)\,d\omega = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^+} \sin\big(\omega(y_t - y_{t-k})\big)g(\omega)\,d\omega \right) dF = 0,$$

where F is the distribution function of y_t. This condition is equivalent to

$$\mathbb{E}[\psi_g(y_t - y_{t-k})] = 0, \tag{5.2}$$

where

$$\psi_g(y_t - y_{t-k}) = \int_{\mathbb{R}^+} \sin\big(\omega(y_t - y_{t-k})\big) g(\omega)\, \mathrm{d}\omega.$$

To test (5.2), Chen, Chou, and Kuan (2000) suggest basing a test on the sample average of $\psi_g(y_t - y_{t-k})$:

$$\mathcal{C}_{g,k} = \sqrt{T_k}\,\bar{\psi}_{g,k}/\bar{\sigma}_{g,k}, \tag{5.3}$$

where $T_k = T - k$ with T the sample size, $\bar{\psi}_{g,k} = \sum_{t=k+1}^{T} \psi_g(y_t - y_{t-k})/T_k$, and $\bar{\sigma}_{g,k}^2$ is a consistent estimator of the asymptotic variance of $\sqrt{T_k}\,\bar{\psi}_{g,k}$. A suitable central limit theorem then ensures that $\mathcal{C}_{g,k}$ is asymptotically distributed as $\mathcal{N}(0,1)$ under the null hypothesis (5.2).

A novel feature of this test is that, because $\psi_g(\cdot)$ is bounded between 1 and -1, *no moment condition is needed* when the central limit theorem is invoked. Yet a major drawback of $\mathcal{C}_{g,k}$ is that the null hypothesis (5.2) is only a necessary condition of (5.1). Indeed, h_k may be integrated to zero by some g function even when h_k is not identically zero. For such a g function, the resulting $\mathcal{C}_{g,k}$ test does not have power against asymmetry of $y_t - y_{t-k}$. Choosing a proper g function is therefore crucial for implementing this test. Chen, Chou, and Kuan (2000) observed that for absolutely continuous distributions, $h_k(\omega)$ is a damped sine wave and eventually decays to zero as $\omega \to \infty$; see Figure 1 for h_k of various "centered" exponential distributions. This suggests choosing g as a function that takes large values for small ω but small values for large ω. A density function of a random variable on $\mathbb{R}^+$ is a potential choice.

Chen, Chou, and Kuan (2000) set g as the density of the exponential distribution ($g = \exp$) with the parameter $\beta > 0$, i.e., $g(\omega) = \exp(-\omega/\beta)/\beta$. This choice of g leads to the following analytic expression for ψ_g:

$$\psi_{\exp}(y_t - y_{t-k}) = \frac{\beta(y_t - y_{t-k})}{1 + \beta^2(y_t - y_{t-k})^2}. \tag{5.4}$$

The closed form (5.4) renders the computation of $\mathcal{C}_{\exp,k}$ test quite easy. One simply plugs the data into (5.4) and calculates their sample average and a consistent estimator for the
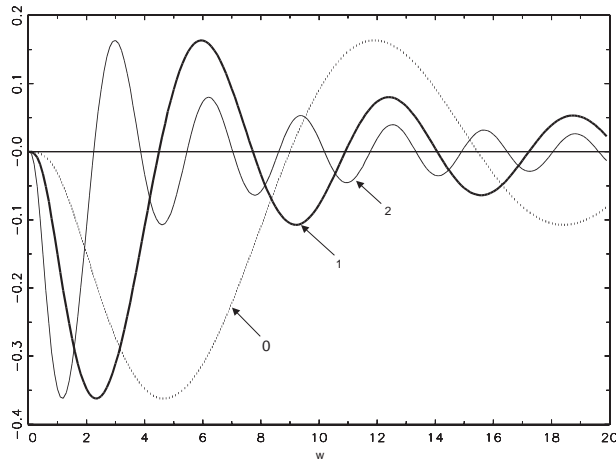
Figure 1: $h(\omega)$ of centered exponential distributions with $\beta = 0.5$ (line 0), $\beta + 1$ (line 1) and $\beta = 2$ (line 2).

standard error. The test statistic is now readily computed as (5.3). It has been shown that the $\mathcal{C}_{\mathrm{exp},k}$ test performs strikingly well in finite samples and is very robust to the data without proper moments. The third-moment-based test and the bi-covariance test, on the other hand, have little power when the data suffer from moment failure.

Chen and Kuan (2002) also demonstrated that the $\mathcal{C}_{\mathrm{exp},k}$ test is very powerful against asymmetric dependence in data. In particular, it is shown that existing tests, such as Q-type tests and the BDS test, fail to distinguish between EGARCH and GARCH models when they are applied to examine the standardized model residuals. Yet, the $\mathcal{C}_{\mathrm{exp},k}$ test is much more sensitive to the difference between symmetry and asymmetry in volatility.

Remarks:

1. One may consider testing a condition equivalent to (5.1). For example, a Cramér-von Mises type condition is based on

$$\int_{\mathbb{R}^+} h_k(\omega)^2 g(\omega) \, \mathrm{d}\omega,$$

   which is zero if, and only if, (5.1) holds. This condition, however, does not permit changing the orders of integration. The resulting test is more difficult to implement and usually has a data-dependent distribution.

2. The $\mathcal{C}_{\mathrm{exp},k}$ test is flexible in practice. By varying the value of β, $\mathcal{C}_{\mathrm{exp},k}$ is able to check departures from (5.2) in different ways. When a small β is chosen, the resulting test

focuses more on $h_k(\omega)$ with smaller ωs. On the other hand, more $h_k(\omega)$ can be taken into account by setting β large. How to choose an optimal β remains an unsolved problem, however.

3. Chen and Kuan (2002) noted that when model residuals are plugged into ψ_g, the resulting standard error in the test (5.3) is more difficult to estimate, due to presence of "estimation effect." A convenient way is to compute a bootstrapped standard error.

References

Bierens, Herman J. (1982). Consistent model specification tests, *Journal of Econometrics*, **20**, 105–134.

Bierens, Herman J. (1984). Model specification testing of time series regressions, *Journal of Econometrics*, **26**, 323–353.

Bierens, Herman J. (1990). A consistent conditional moment test of functional forms, *Econometrica*, **58**, 1443–1458.

Bierens, Herman J. and Werner Ploberger (1997). Asymptotic theory of integrated conditional moment tests, *Econometrica*, **65**, 1129–1151.

Box, G. E. P. and David A. Pierce (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American Statistical Association*, **65**, 1509–1526.

Brock, W. A., W. D. Dechert and J. A. Scheinkman (1987). A test for independence based on the correlation dimension, Working paper, Department of Economics, University of Wisconsin at Madison.

Brock, W. A., W. D. Dechert, J. A. Scheinkman, and B. LeBaron (1996). A test for independence based on the correlation dimension, *Econometric Reviews*, **15**, 197–235.

Brooks, Chris and Saeed M. Heravi (1999). The effect of (mis-specified) GARCH filters on the finite sample distribution of the BDS test, *Computational Economics*, **13**, 147–162.

Brooks, Chris and Óland T. Henry (2000). Can portmanteau nonlinearity tests serve as general mis-specification tests? Evidence from symmetric and asymmetric GARCH models, *Economics Letters*, **67**, 245–251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

Chen, Yi-Ting, Ray Chou, and Chung-Ming Kuan (2000). Testing time reversibility without moment restrictions, *Journal of Econometrics*, **95**, 199–218.

Chen, Yi-Ting and Chung-Ming Kuan (2002). Time irreversibility and EGARCH effects in U.S. stock index returns, *Journal of Applied Econometrics*, **17**, 565–578.

Cochrane, John H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, **96**, 893–920.

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments, *Scandinavian Journal of Statistics*, **8**, 93–115.

de Lima, Pedro J. F. (1997). On the robustness of nonlinearity tests to moment condition failure, *Journal of Econometrics*, **76**, 251–280.

Deo, Rohit S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity, *Journal of Econometrics*, **99**, 291–315.

Dominguez, Manuel A. and Ignacio N. Lobato (2000). A consistent test for the martingale difference hypothesis, Working Paper, Instituto Tecnologico Autonomo de Mexico.

Durlauf, Steven N. (1991). Spectral based testing of the martingale hypothesis, *Journal of Econometrics*, **50**, 355–376.

Fuller, Wayne A. (1976). *Introduction to Statistical Time Series*, New York, NY: Wiley.

Hausman, Jerry A. (1978). Specification tests in econometrics, *Econometrica*, **46**, 1251–1271.

Hong, Yongmiao (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach, *Journal of the American Statistical Association*, **94**, 1201–1220.

Hsieh, David A. (1989). Testing for nonlinear dependence in daily foreign exchange rates, *Journal of Business*, **62**, 339–368.

Hsieh, David A. (1991). Chaos and nonlinear dynamics: Application to financial markets, *Journal of Finance*, **46**, 1839–1877.

Hsieh, David A. (1993). Implications of nonlinear dynamics for financial risk management, *Journal of Financial and Quantitative Analysis*, **28**, 41–64.

Kuan, Chung-Ming and Wei-Ming Lee (2004). A new test for the martingale difference hypothesis, *Studies in Nonlinear Dynamics and Econometrics*, **8:4**, Article 1.

Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models, *Biometrika*, **65**, 297–303.

Lobato, Ignacio, John C. Nankervis, and N. E. Savin (2001). Testing for autocorrelation using a modified Box-Pierce Q test, *International Economic Review*, **42**, 187–205.

McLeod, A. I. and W. K. Li (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations, *Journal of Time Series Analysis*, **4**, 269–273.

Park, Joon Y. and Yoon-Jae Whang (2005). A test of the martingale hypothesis, *Studies on Nonlinear Dynamics and Econometrics*, **9:2**, Article 2.

Pinkse, Joris (1998). A consistent nonparametric test for serial independence, *Journal of Econometrics*, **84**, 205–231.

Ramsey, James B. and Philip Rothman (1996). Time irreversibility and business cycle asymmetry, *Journal of Money, Credit, and Banking*, **28**, 3–20.

Rothman, Philip (1992). The comparative power of the tr test against simple threshold models, *Journal of Applied Econometrics*, **7**, S187–S195.

Shorack, Galen R. and John A. Wellner (1986). *Empirical Processes with Applications to Statistics*, New York, NY: Wiley.

Skaug, Hans J. and Dag Tjostheim (1993). A nonparametric test of serial independence based on the empirical distribution function, *Biometrika*, **80**, 591–602.

Taylor, Stephen J. (2005). *Asset Price Dynamics, Volatility, and Prediction*, Princeton, NJ: Princeton University Press.

Tong, Howell (1990). *Non-linear Time Series, A Dynamic System Approach*, New York, NY: Oxford University Press.

Whang, Yoon-Jae (2000). Consistent bootstrap tests of parametric regression functions, *Journal of Econometrics*, **98**, 27–46.

Whang, Yoon-Jae (2001). Consistent specification testing for conditional moment restrictions, *Economics Letters*, **71**, 299–306.